

Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis

Year: 2020 | Citations: 540 | Authors: Agostina J. Larrazabal, Nicolás Nieto, Victoria Peterson, Diego H. Milone, Enzo Ferrante

Abstract

Artificial intelligence (AI) systems for computer-aided diagnosis and image-based screening are being adopted worldwide by medical institutions. In such a context, generating fair and unbiased classifiers becomes of paramount importance. The research community of medical image computing is making great efforts in developing more accurate algorithms to assist medical doctors in the difficult task of disease diagnosis. However, little attention is paid to the way databases are collected and how this may influence the performance of AI systems. Our study sheds light on the importance of gender balance in medical imaging datasets used to train AI systems for computer-assisted diagnosis. We provide empirical evidence supported by a large-scale study, based on three deep neural network architectures and two well-known publicly available X-ray image datasets used to diagnose various thoracic diseases under different gender imbalance conditions. We found a consistent decrease in performance for underrepresented genders when a minimum balance is not fulfilled. This raises the alarm for national agencies in charge of regulating and approving computer-assisted diagnosis systems, which should include explicit gender balance and diversity recommendations. We also establish an open problem for the academic medical image computing community which needs to be addressed by novel algorithms endowed with robustness to gender imbalance.